This is a take home final exam. You can use your notes, my online notes on canvas and the text book. You are supposed to work on your own text without external help. I'll be available to answer question in person or via email. Please, write clearly and legibly and take a readable scan before uploading.

To solve the Exam problems, I have not collaborated with anyone nor sought external help and the material presented is the result of my own work.

Signature: ______________________

Name: ______________________

Question:	1	2	3	Total
Points:	35	25	40	100
Score:				

Question:	1	2	3	Total
Bonus Points:	0	0	30	30
Score:				

Question 1 .... *35 point*

Let x_i, $i = 1, \ldots, 48$, be the result of 48 independent measurements of the amount of pollen grains in the air in Atlanta in the spring of 2020. You know that

$$\sum_{i=1}^{48} x_i = 492.02 \qquad \sum_{i=1}^{48} x_i^2 = 5420.4$$

You can use both Matlab and R (or any other software) to compute the needed values of the c.d.f. and inverse c.d.f. of a central or non central t distribution (see page 579 in the textbook). In case you do not have them available, this is an online t-distribution calculator.

(a) (10 points) Assume that the results come from a normal distribution with both expected value μ and variance σ^2 unknown. Find the coefficient 0.99 Confidence Interval for μ.

Solution: We have

$$\bar{x} = \frac{492.02}{48} = 10.25 \qquad s^2 = \frac{1}{47}\left(5420.4 - \frac{492.02^2}{48}\right) = 8.02$$

so that, using that $t_{47,0.005} = 2.68$ we get

$$10.25 - \frac{2.68\sqrt{8.02}}{\sqrt{48}} \leq \mu \leq 10.25 + \frac{2.68\sqrt{8.02}}{\sqrt{48}}$$

or

$$9.15 \leq \mu \leq 11.35$$

(b) (15 points) You decide to hold the following test for the true population mean μ

- H_0: $\mu \geq \mu_0$ where $\mu_0 = 11$
- H_a: $\mu < \mu_0$

at a significance level of 0.01. Give rejection region for the test of size 0.01. Will you reject H_0 at this significance level?

Compute the p-value of the above test.

Solution: If

$$T = \frac{\sqrt{N}(\bar{X} - \mu_0)}{\sqrt{\dfrac{S^2}{N-1}}}$$

then the test is $T < -t_{47,0.01} = -2.40$. With our data we get

$$t = -1.83$$

so that we do not reject H_0.

The p-value of the test is $\Pr(T < -1.83) = 0.037$ where T is a t r.v. with 47 d.o.f. Thus

$$P = 0.037\,.$$

(c) (10 points) Compute the probability $\beta(10, 8)$ of a Type II error when the true expected value is $\mu = 10$ and the true value of $\sigma^2 = 8$.

Suppose you want a test of the null hypothesis H_0 of point b) with size 0.01 and $\beta(10, 8) < 0.2$. Give an estimate or lower bound of the sample size you will need.

Solution: Calling $T_N(x|\psi) = \Pr(T \leq x)$ when T is a noncentral t r.v. with N d.o.f. and noncentrality ψ we have, with $\psi = -\sqrt{48/8} = -2.45$,

$$\beta(10, 8) = 1 - T_{47}(-2.40| - 2.45) = 0.48$$

We need N such that

$$\beta(10, 8) = 1 - T_{N-1}\left(T_{N-1}^{-1}(0.01\,|\,0)\,\middle|\, - \sqrt{\frac{N}{8}}\right) < 0.2\,.$$

The smallest such N is $N = 84$.

Question 2 . . . . *25 point*

In some problem of anomalous conduction the distribution of the speed V (the modulus of the velocity) of a particle is given by

$$f(v \mid \mu) = \begin{cases} \frac{\mu^{\frac{1}{3}}}{\Gamma(4/3)} e^{-\mu v^3} & v \geq 0 \\ \\ 0 & \text{otherwise} \end{cases}$$

Suppose you ran a sample $\mathbf{V}$ of size N from the distribution $f(v \mid \mu)$ and obtained $\mathbf{v}$ as a result.

(a) (10 points) Show that the $\Gamma(\alpha, \beta)$ family of distributions for μ forms a conjugate family of priors for samples from the distribution $f(v \mid \mu)$. Compute the hyperparameter $\alpha_1(\mathbf{v})$ and $\beta_1(\mathbf{v})$ of the posterior distribution if the prior distribution is $\Gamma(\alpha_0, \beta_0)$. Assuming a square error loss function, find the Bayes estimator $\hat{\mu}_B(\mathbf{V})$.

Solution: If the prior distribution $g(\mu)$ for μ is $\Gamma(\alpha_0, \beta_0)$ we get that the posterior distribution is

$$g(\mu \mid \mathbf{v}) \propto \mu^{\alpha_0 - 1} e^{\beta_0 \mu} \mu^{\frac{1}{3}N} e^{-\mu \sum_{i=1}^N v_i^3}$$

so that the hyperparameter of the posterior distribution are:

$$\beta_1(\mathbf{v}) = \beta_0 + \sum_{i=1}^{N} v_i^3$$
$$\alpha_1(\mathbf{v}) = \alpha_0 + \frac{1}{3}N$$

For a square error loss function the Bayes estimator is the mean of the posterior distribution, that is

$$\hat{\mu}(\mathbf{V}) = \frac{\alpha_0 + \frac{1}{3}N}{\beta_0 + \sum_{i=1}^N v_i^3} .$$

(b) (15 points) Find the Maximum Likelihood estimator $\hat{\mu}_{MLE}(\mathbf{V})$ for μ and discuss its asymptotic distribution for large N.

Solution:
The likelihood function is

$$l(\mu) \propto \mu^{\frac{1}{3}N} e^{\mu \sum_{i=1}^{N} v_i^3} \quad \Rightarrow \quad l'(\mu) \propto \frac{N}{3} \mu^{\frac{1}{3}N-1} e^{\mu \sum_{i=1}^{N} v_i^3} - \mu^{\frac{1}{3}N} e^{\mu \sum_{i=1}^{N} v_i^3} \sum_{i=1}^{N} v_i^3$$

so that we find

$$l'(\mu) = 0 \quad \Rightarrow \quad \mu = \frac{3 \sum_{i=1}^{N} v_i^3}{N}$$

Since $l(0) = l(\infty) = 0$ and the function has a unique critical point, it must be a maximum and we have:

$$\hat{\mu}_{MLE}(\mathbf{V}) = \frac{N}{3 \sum_{i=1}^{N} V_i^3}.$$

Asymptotically $\hat{\mu}_{MLE}(\mathbf{V})$ has a Normal distribution with mean μ and variance σ^2 given by

$$\sigma^2 = \frac{1}{-N\mathbb{E}(\partial_\mu^2 \ln f(V \mid \mu))} = \frac{3\mu^2}{N}\,.$$

Question 3 . . . . *40 point*

A random sample $\mathbf{X}$ of size N is selected from a population with distribution

$$f(x\,|\,\theta, A) = \begin{cases} \frac{1}{\theta}e^{-\frac{(x-A)}{\theta}} & x \geq A \\ \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where $\theta > 0$ and A are parameters.

(a) (10 points) Use Maximum Likelihood to find estimators $\hat{\theta}(\mathbf{X})$ for θ and $\hat{A}(\mathbf{X})$ for A.

Solution: Let X_i be the r.v. forming your sample. The likelihood function is

$$l(A, \theta) = \theta^{-N} \exp\left(-\sum_{i=1}^{N} \frac{(x_i - A)}{\theta}\right)$$

if $x_i > A$ for every i and 0 otherwise. Thus the maximum over A is reached when A is equal to the smallest of the x_i that is

$$\hat{A} = \min_i X_i := \underline{X}$$

fixed A we can differentiate over θ and get

$$\hat{\theta} = \frac{1}{N}\sum_{i=1}^{N} X_i - \hat{A} = \overline{X} - \underline{X}$$

(b) (15 points) Assuming that θ is known, you want to test the hypotheses

- H_0: $A \leq A_0$
- H_a: $A > A_0$

You decide to use the statistics

$$R = \frac{\hat{A} - A_0}{\theta/N}$$

and the test δ_c that rejects H_0 if $R > c$. Compute the power function $\pi(A|\delta_c)$ and use it to find c such that the test δ_c has size α. (**Hint:** find first the p.d.f. of $\hat{A}$.)

Solution:

Since the X_i are independent we have

$$\Pr(\hat{A} > a) = \prod_i \Pr(\hat{X}_i > a) = \begin{cases} \exp\left(-\frac{N(a-A)}{\theta}\right) & a > A \\ 0 & \text{otherwise} \end{cases}$$

Clearly a value of $\hat{A}$ much bigger than A_0 would be a strong evidence against H_0. This means a value of

$$R = \frac{\hat{A} - A_0}{\theta/N}$$

greater than 0. This is why our test is $\delta : R > c$. Observe that $R - N(A - A_0)/\theta$ is a exponential r.v. with mean 1. Thus the power function is

$$\pi(\delta_c|A) = \Pr(R > c \,|\, A) = \begin{cases} e^{N(A-A_0)/\theta - c} & A < A_0 - \theta c/N \\ 1 & \text{otherwise} \end{cases}$$

$$\sup_{A<A_0} \pi(\delta_c|A) = \sup_{A<A_0} e^{N(A-A_0)/\theta - c} = e^{-c}$$

so that $c = -\log(\alpha)$ and

$$\delta_\alpha : \; R > -\log(\alpha)\,.$$

(c) (15 points) Find the p-value (as a function of the sample result) for the test in part (c).

Solution:

If r is the result of the sample for the R statistics we get

$$P = \sup_{A<A_0} \Pr(R > r \mid A) = e^{-r}$$

(d) (15 points (bonus)) Assume now that θ is unknown, find the joint p.d.f. of $\hat{\theta}$ and $\hat{A}$. (**Hint:** if X_1 and X_2 are two independent exponential r.v. with the same expected value, find the j.p.d.f. of $Y_1 = \min(X_1, X_2)$ and $Y_2 = X_1 + X_2$. A way to do this is to look at $\Pr(Y_2 < t \,|\, Y_1 > s)$.)

Solution: We know that

$$\hat{A} = \min_i X_i \qquad\qquad N\hat{\theta} = \sum_i X_i - N \min_i X_i$$

so that we start computing

$$\Pr\left(\sum_i X_i - N\min_i X_i > c \,\&\, \min_i X_i > d\right)$$

for $c > 0$ and $d > A$. Since $\Pr(X_i = X_j) = 0$ we can write

$$\Pr\left(\sum_i X_i - N\min_i X_i > c \,\&\, \min_i X_i > d\right) =$$
$$N\Pr\left(\sum_i X_i - N\min_i X_i > c \,\&\, \min_i X_i > d \,\&\, \operatorname{argmin}_i X_i = 1\right) =$$
$$N\Pr\left(\sum_{i=2}^N (X_i - X_1) > c \,\&\, X_i - X_1 > 0\,, \forall i > 2 \,\&\, X_1 > d\right) =$$
$$\frac{N}{\theta^N}\int_d^\infty dx_1 \int_{\substack{x_i - x_1 > 0 \\ \sum_{i=2}^N (x_i - x_1) > c}} e^{-\sum_i (x_i - A)/\theta} dx_2 \cdots dx_N =$$
$$\frac{N}{\theta^N}\int_d^\infty dx_1 e^{-N(x_1 - A)/\theta} \int_{\substack{x_i - x_1 > 0 \\ \sum_{i=2}^N (x_i - x_1) > c}} e^{-\sum_{i=2}^N (x_i - x_1)/\theta} dx_2 \cdots dx_N \qquad y_i = \frac{x_i - x_1}{\theta}$$
$$\frac{N}{\theta}\int_d^\infty dx_1 e^{-N(x_1 - A)/\theta} \int_{\substack{y_i > 0 \\ \sum_{i=2}^N y_i > c/\theta}} e^{-\sum_{i=2}^N y_i} dy_2 \cdots dy_N =$$
$$e^{-N(d-A)/\theta}(1 - \Gamma_{N-1,1}(c/\theta))\,.$$

It follows that $\hat{\theta}$ and $\hat{A}$ are independent and $N(\hat{A} - A)/\theta$ is exponential with expected value 1 while $N\hat{\theta}/\theta$ is Gamma with $\alpha = N - 1$ and $\beta = 1$.

(e) (15 points (bonus)) Still assuming that θ is unknown, use the result of part (d) to find a test δ for the hypotheses in part (b) at a given level α. Discuss the power function and p-value of this test.

Solution: Consider the statistics

$$S = \frac{\hat{A} - A_0}{\hat{\theta}} = \frac{N(\hat{A} - A_0)/\theta}{N\hat{\theta}/\theta} = S' + \frac{A - A_0}{\hat{\theta}}$$

and the rejection region $S > c$. Observe that $S' = W/U$ where W is exponential with average 1 and u is Gamma with $\alpha = N - 1$ and $\beta = 1$. Thus we have

$$\Pr(S > c|A_0, \theta) = \int_0^\infty \Pr(W > cu \,|\, U = u) f_U(u) du =$$
$$\int_0^\infty e^{-cu} \frac{u^{N-2}}{\Gamma(N-1)} e^{-u} du = \frac{1}{(1+c)^{N-1}}$$

moreover if $A < A_0$ we have

$$\Pr(S > c|A, \theta) = \Pr(S' > c + (A_0 - A)/\hat{\theta}|A, \theta) < \Pr(S' > c|A) = \Pr(S > c|A_0, \theta)$$

so that

$$\pi(A, \theta \,|\, \delta_c) < \pi(A_0, \theta \,|\, \delta_c) \quad if \quad A < A_0$$
$$\pi(A, \theta \,|\, \delta_c) > \pi(A_0, \theta \,|\, \delta_c) \quad if \quad A > A_0$$

and

$$\sup_{A < A_0} \pi(A, \theta \,|\, \delta_c) = \pi(A_0, \theta \,|\, \delta_c) =$$
$$= \Pr(S > c \,|\, A_0, \theta) = \frac{1}{(1+c)^{N-1}}\,.$$

For the test to have size α we need $c = \alpha^{-1/N} - 1$. In the same way if the sample give a value s for the statistics S the p-value is $(1+s)^{-(N-1)}$.
Finally the p.d.f. of S for every A depends only on $N(A - A_0)/\theta$ as a kind of non-centrality parameter.